# NEW TRENDS IN SOUND SYNTHESIS AND AUTOMATIC TUNING OF ELECTRONIC MUSICAL INSTRUMENTS

**Juan R. Aguilar** [1]    **Renato Salinas** [2]

## *ABSTRACT*

This paper is divided in two parts: first part corresponds to a review of newest sound synthesis techniques, while the second part contains the development of an automatic tuning system based on a fuzzy logic controller intended to emulate the human expertise in this field. As part of this work, two new concepts will be introduced: first, the idea of "tuning" a synthesized musical instrument, and second, the application of automatic tuning procedures by means of intelligent systems. If we consider the synthesis as a whole process which involves the reproduction of every aspect that determines the realistic sound of the musical instruments, tuning the synthesizer as a human expert does becomes meaningful.

*Keywords: Sound Synthesis, Automatic Tuning, Musical Instruments.*

## INTRODUCTION

Traditional sound synthesis techniques are those based on Fourier theory. Starting from a filter bank, and by means of additive or subtractive synthesis techniques, any spectrum can be reconstructed. However, the arrival of new models representing the "naturalness" of vibration in musical instruments, lead to new techniques in sound synthesis.

On the other hand, as is it well known, tuning of musical instruments involve matching the frequency of notes in such a way to achieve harmonic relations between them. In particular, tuning of large keyboard instruments has been always problematic due to multiples scales. The case of plucked or struck string keyboard instruments represent a special situation where physical behavior of strings lead to additional difficulties.

As an approach to automatic piano tuning, we propose a fuzzy logic based system were we capture the know-how of an experienced piano tuner in the form of rules and procedures.

We implemented our system, tested it, and from the results, we suggest ideas for further researches.

## THEORETICAL BACKGROUND

### Sound Synthesis by Numerical Simulation of Differential Equations

As an alternative and more interesting approach to sound synthesis, it has been proposed the numerical solution of differential equations.

For example, in the case of piano string vibration we have the differential equation:

$$\frac{\partial^2 y}{\partial t^2} = c^2\frac{\partial^2 y}{\partial x^2} - \varepsilon c^2 L^2 \frac{\partial^4 y}{\partial x^4} - 2b_1\frac{\partial y}{\partial t} + 2b_3\frac{\partial^3 y}{\partial t^3} + f(x,y,t) \quad (1)$$

which includes terms for frequency dependent loss factor $b_1$ and $b_3$, stiffness $\varepsilon$ for bending wave behavior and a $f(x,y,t)$ force term for nonlinear hammerexcitation. To solve this equation it will be necessary to specify initial and boundary conditions [1].
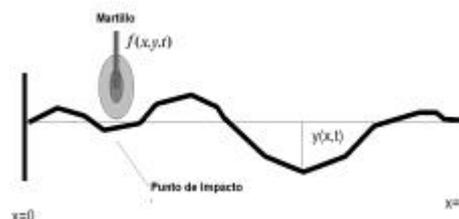


Fig.1.-    Sketch of the discrete piano string and hammer

[1]  University of Santiago of Chile, Acoustical Engineer Ph.D. on Automation Program. Electrical Engineering, Department, Ave. Ecuador 3519. Santiago. Chile, Chileacus@hotmail.com.

[2]  University of Santiago of Chile, Electrical Engineering Department, Ave. Ecuador 3519. Santiago. Chile. Rsalinas@lauca.usach.cl.

The simplest way to solve the equation (1) is by a system of difference equations which transform the continuous wave equation into a fourth order finite difference scheme, both in spatial and time domains, giving the discrete string sketch of Fig. 1, and the following recursive algorithm:

$$
\begin{aligned}
y(x,(t+1)) = & \, a_1 y(x,t) + a_2 y(x,(t-1)) \\
& + a_3[y((x+1),t) + y((x-1),t))] \\
& + a_4[y((x+2),t) + y((x-2),t))] \\
& + a_5 \begin{bmatrix} y((x+1),(t-1)) + y((x-1),(t-1)) \\ + y(x,(t-2)) \end{bmatrix} \\
& + hammer \ force \ term
\end{aligned}
\tag{2}
$$

Where the numerical stability of the model is given by [3].

$$
N_{MAX} = \sqrt{\frac{-1 + \sqrt{1 + 16\varepsilon\left(\frac{f_S}{2f_1}\right)^2}}{8\varepsilon}}
\tag{3}
$$

where $N_{MAX}$ is the maximum number of length steps or intervals, $\varepsilon$ is the stiffness parameter, $f_S$ is the sampling frequency and $f_1$ is the fundamental frequency of vibration.

Clearly, this kind of numerical simulation of piano string vibration does not take into account acoustical radiation from soundboard of the instrument nor bending wave acoustical behavior of wave coincidence phenomenon.

## SOUND SYNTHESIS BY NEURAL NETWORK BASED LINEAR PREDICTION

Neural network are mathematical structures characterized by their similarity with real human neurons in terms of their synaptic connection capabilities. There are plenty of models of neurons. For linear systems prediction, we require linear neurons with supervised learning. This way, it is possible to train the network from N-input vector data; N being the same as the number of delay units of the input signal.

The original time-series corresponds to an audio signal captured while tuning a piano. This signal was generated by playing, two keys simultaneously. The frequency ratio between these two notes is 4:3.

Audio data, in the form of a vector, without delays, was required to correspond to the input signal. That is to say,

if the input to the neuron considers the 10 initial elements from the input vector, we shall train the neural net to predict the following element in the sequence, i.e., the 11th. Even though we can potentially generate long sequences by repeating the procedure above, our experiments have shown that the amount of computational resources require to simulate the neural net, slows downs the whole system.

## TUNING AND TEMPERAMENT OF MUSICAL INSTRUMENT

The definition of the "interval" between two notes, for example, the interval between C and A, is called a "sixth" because there are six notes between them. Pythagoras was the first person who defined the mathematical frequency relationships inside an octave that results in the pure intervals.

Table N°1.- Pythagorean frequency ratios for music intervals

| Interval | Frequency Ratio |
|----------|-----------------|
| Octave | 2:1 |
| Major Sixth | 5:3 |
| Fifth | 3:2 |
| Fourth | 4:3 |
| Major Third | 5:4 |
| Minor Third | 6:5 |

The relation shown in preceding table holds for perfectly harmonic interval or pure intervals. As a consequence of these ratios, it results the perfect matching between components of the spectra of the two notes. For example, in the case of a Fourth A – D (4:3), Fig. 2, it means that the fourth harmonic of the lower note A must coincide with the third harmonic of the higher note D, to reach the perfect tuning.
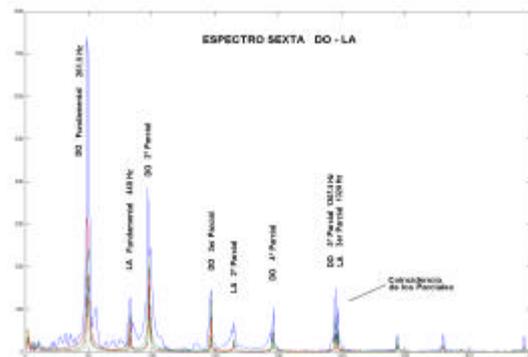


Fig.2.- Frequency spectrum of Do-La interval

However, these perfect coincidences are correct only inside an octave, because for intervals which comprise more than one octave the Pythagorean relationships do

not hold exactly. For example, let us consider a harmonic series of 6 Major Sixth, which theoretically must be equal to a 3 Octave Series, however, it is possible to observe some discrepancies,

$$6 \text{ Maj. Sixth Series: } (5/3)^6 = 21.43$$
$$3 \text{ Octave Series: } (2)^3 = 8.00$$
$$\text{Difference: } 13.43$$

this frequency mismatch can be observed in the form of temporal amplitude modulation at near coincident harmonics, producing *beats*.

These discrepancies, were early observed by Pythagoras, and vary depending upon the interval considered. So, it is evident, that the beat frequency in relation with the kind of interval being tuned, is the basic parameter that the expert uses to tune a piano.

In the case of the piano, and other large keyboard instruments comprising a large number octaves, it has became usual practice to accept compromises in tuning intervals, stretching a little the Fifths and enlarging a little the Fourths, these compromises are called the "Temperament" of the scale. Accepting such compromises in tuning has become necessary and the way of understanding these compromises has been changing throughout the evolution of occidental music, producing at least four eras: Pythagorean Tuning, Meantone Tuning, Well Tempered Tuning, and Equally Tempered Tuning [2], [3].

## INTELLIGENT SYSTEM FOR THE AUTOMATIC CONTROL OF TUNING

Fuzzy logic has demonstrated to be well suited in controlling complex processes, and/or in processes where it is typically difficult to develop a model. A fuzzy system, basically, comprises a fuzzyfier, a controller and a defuzzyfier, [4].

The fuzzifier is defined by a set of membership functions used to specify the range over which a crisp input can vary. The membership functions are named as linguistic variables, for example: high, near, low, etc. The correspondence of the input with the membership functions represents the way that the fuzzyfier converts input data into fuzzy data.

The controller is defined by a set of user fuzzy if-then rules, such as:

*if* **beat frequency** *is* **low** *then* **string tension** *is* **go_up**
*if* **beat frequency** *is* **near** *then* **string tension** *is* **tuned**

and so on. These fuzzy implications describe how the system is being controlled. After the fuzzy rules have been executed (or fired), the defuzzyfier, which is also defined by another set of output memberships functions, assigns concrete values to the output of the fuzzy controller Neuro fuzzy systems have been used for adaptive control of musical processes and for automatic classification of musical styles, [5], but our literature survey has not found references on early applications of tuning of musical instrument by using fuzzy logic systems.

As shown in an earlier section, piano tuning procedures are particularly fuzzy, both in the sense of the physical phenomenon involved as well as in the historical evolution of tuning criteria.

The input variable in our fuzzy controller will be, most naturally, the beats of detune between any partials corresponding to each interval being tuned. Consequently, the output variable in our fuzzy controller, will be the tension T of the synthetic string which is being tuned.

## EXPERIMENTAL PROCEDURES

### Piano Synthesis by Numerical Solution of Vibrating String Differential Equation

The mathematical expression of Eq. 2 was computationally implemented. The application of this recursion formula results in a discrete string. The result of this simulation was an N x M matrix, where N represents the number of discrete string length segments and M, the number of time steps of the whole simulation event.

A graph showing the first 10 mSec of a 100 segment discrete piano string is illustrated in Fig. 3.- . It is easy to observe how the perturbation generated by the hammer propagates along the string and reflects at its boundaries.

For signal analysis purposes, only one row of the matrix was used; this is equivalent to pick up the vibration at one point on the string. The simulation time was 1 second, with 20 string segments and a sampling frequency of 44.1 KHz. Using a Pentium processor at 500 MHz almost 3 hours were required by the computer to perform the calculations.

Fig. 4.- (upper graph) shows the frequency spectrum of the original recorded signal, compared with the spectrum of the synthetic signal (lower graph). The musical interval corresponds to a Fourth with the lower

note D tuned near 293.665 Hz and the higher note A at 440 Hz. In the spectra it is possible to observe the coincidence at 4:3 partials, as well as frequency shifting due to dispersion phenomena.
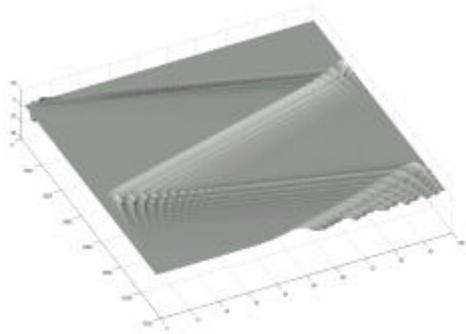


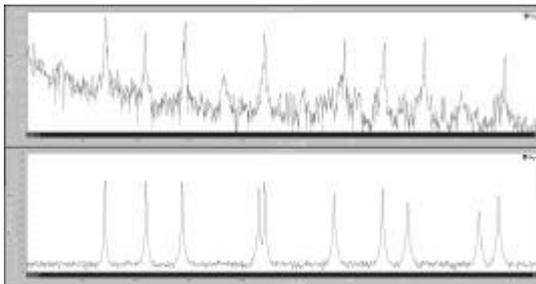Fig. 3.- Graphical representatio n of the first 10 mSec of simulated piano string vibration.



Fig. 4.- Comparison of real and synthetic spectrums of Fourth D-A interval.

Fig. 5.- shows the time-frequency spectrum of a synthetic sixth interval. First 50 components of a 512 point STFFT, here it is possible to appreciate and measure the beat at near coincident 4:3 partials, as well as frequency dependent damping of synthetic notes.
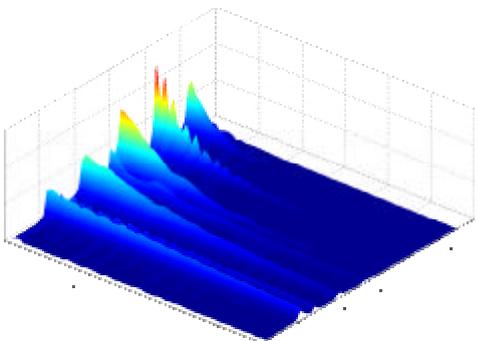


Fig. 5.- Short Time Fast Fourier Transform depicting of Sixth interval

**Piano Synthesis by Neural Network based Linear Prediction**

Due to the fact that present and future values of the input signal are either stable or slowly varying, the neural net can be trained from input-output examples. It has been proposed a Back Propagation (BP) net with 5 input nodes, 5 hidden nodes and 1 output node. All transfer functions were chosen as pure linear.

We selected the Levenberg-Marquardt minimization criteria to train the net and a temporal sequence corresponding to a recorded sample of two notes Mi and Re was used as training data set.

The net was trained near 10 epochs at a training error goal of $10^{-4}$.

### SIMULATION EXPERIMENTS

The (linear) neural network, after training, was then tested with new data from a signal of similar characteristics as the employed for training. This new signal was also captured during our sound recording sessions.

This new signal also corresponds to the same notes Mi and Re played together.

*Matlab code is as follow:*

```
dato = wavread('cuarta2');
timeseries = dato(1:22050)';
L = length(timeseries);


Q = zeros(1,L);
Q(1,1:11) = timeseries(1,11);
for q = 12:L
  li=q-11;
  ls=q-2;
  T = P(1,li:ls);
  Q(1,q) = sim(net,T');
end

wavwrite(Q,44100,'neuro_audio2b');
```

To validate the performance of our neuronal model, we need to compute the error.

error = Q2 - timeseries

Since the signals do not exactly coincide time-wise, it was observed a time-shift of about 20 samples. To

compute the error vector, it is necessary to correlate these signal. We chose to make the maxima coincide.

After plotting this error signal, it is easy to see that the performance of our NN trained as a sound synthesizer, was quite remarkable, with average values near $10^{-3}$. (Fig. 6)
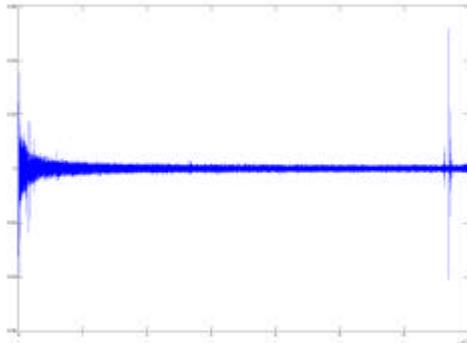


Fig. 6.-   Neural network simulation error

## AUTOMATIC PIANO TUNING SYSTEM

### Tuning Session Recording

Experimental procedures began with the recording of sound during a real tuning session, because it was useful to know the spectral characteristics of the tuning events to compare those characteristics with the synthesized signal. The recording was done using a "hard disk" system, shown in Fig. 7.



Fig. 7.-   The microphone and hard disk recording setup

Recording was done at 24 bits of resolution and 48 KHz of sampling frequency. With this system, 16 bits .wav files were available for computational treatment of the data. A Type 1 sound level meter microphone was used with ⅞ diameter, random incidence calibration and low sensibility. The microphone was mounted in a shock absorber at about 20 cm from the string. A complete tuning session of a vertical piano was then recorded .

## TIME-FREQUENCY ANALYSIS OF RECORDINGS AND BEAT MEASUREMENT

Using Matlab, Short Time Fast Fourier Transform (STFFT) analysis was applied to the recorded signal. This technique of frequency analysis enables us to extract temporal modulation of amplitude at coincident spectral components when two notes of the piano are played together.

The beat for the case of a sixth can be appreciated in the Fig. 8. Then, a computational routine was implemented to isolate the coincident spectral components and measure periodicity of the beat, as shown in Fig. 9. This variable will correspond to the input of our control system.
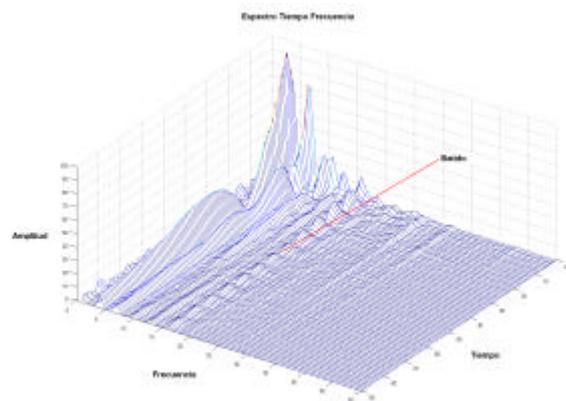


Fig. 8.-   Time-Frequency spectrum of recorded C-A Sixth interval, revealing beat at near coincident partials.
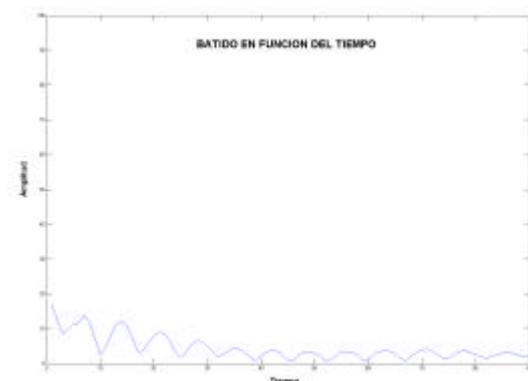


Fig. 9.-   .Beating decay of coincident spectral components. Beat frequency =  11.87 Hz.

## FUZZY LOGIC BASED CONTROL SYSTEM

Three Gaussian input membership functions were specified to fuzzify the beat frequency variable. The

functions were associated to the linguistic variables: low, near and high, see Fig. 10.

A fuzzy inference system composed of three rules was defined, and the rules were:

*1. If beat is **low** then output  is  **high***
*2. If beat is **near** then output  is  **in tune***
*3. If beat is **high** then output   is  **low***

These rules describe how tuning is being controlled.

The centroid operator was selected to defuzzify, and trapezoidal and triangular output membership functions were specified. The functions were associated  with the linguistic variables:  *go down*, *go_up*, and *tuned*, describing the form in which the string tension must beadjusted before the next simulation cycle.
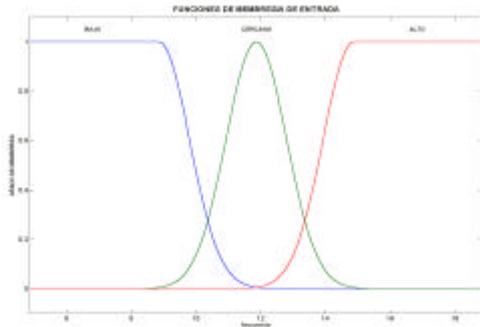


Fig. 10.-    Gaussian input membership functions of the fuzzy controller
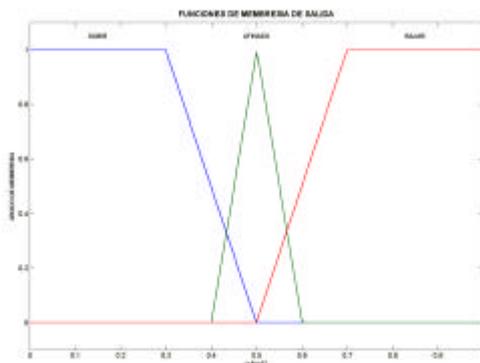


Fig. 11.-    Trapezoidal and triangular output membership functions of the fuzzy controller

## CONCLUSIONS

Conclusions for this work consider two major aspects:

Regarding to the sound synthesis, the differential equation approach seems to be efficient but is limited because it does not incorporate acoustical radiation effects, such as resonances from the piano soundboard and wave coincidence phenomena of bending wave propagation along the string. On the other hand, neural network based linear prediction synthesis should be called "sampler" synthesis.

Second, we believe that this paper represents a first attempt towards the goal of automatic piano tuning.

Automatic tuning could become a new and powerful approach that provides an objective procedure for this task that is otherwise highly subjective.

An important result of this work could be that it would be feasible to have a permanent tuning system working continuously to ensure the proper tuning of the instrument.

A possible limitation of the current version might be that is built around just only one expert. Therefore, the fuzzy inferences reflect his or her particular tuning criteria.

We think that further research is needed to test other tuning techniques based on different intelligent approaches, such as neural networks, genetic algorithms, and so forth.

Future developments could include new and improved membership functions, additional rules, as well as hardware implementation of the string simulation equations and the fuzzy controller. We could also imagine the day in which a piano tuner could remotely tune the synthesizer via Internet, or he/she can perform on-line monitoring of  the tuning status of the electronic instrument.

## REFERENCES

[1]  A. Chaigne and A. Askenfelt; "Numerical simulation of piano strings. I. A physical model for a struck string using finite difference methods".

Journal of the Acoustical Society of America, Vol. 95 (2), February, 1994.

[2] A. Reblitz; "Piano servicing, tuning and rebuilding". The Vestal Press Ltd. New York, 1993.

[3] Edward Foote; "Historical tunings on the modern concert Grand Piano". http://www.uk-piano.org/edfoote/.

[4] Lefteri Tsoukalas, Robert Uhrigh; "Fuzzy and neural approaches in engineering". John Wiley and Sons, 1997.

[5] Michael Lee and David Wessel; "Real-Time Neuro-Fuzzy Systems for Adaptive Control of Musical Processes". ICMC. 1993.